

## StudiesReallyMeanWhat?

Submitted by drhall on July 26, 2005 - 8:05am.

TESTING NOTES BearCreek Research - Not all studies are definitive, In fact, very few are.

Science is a BEST GUESS estimate, with one study building on the last until some clear proof emerges over time.

Bear Creek is more involved with Meta-Analysis than the other types.

TYPES OF STUDY Be It a Cohort Study (Beginning with a group of people who have no disease and following), Epidemiological studies. These are studies looking for potential disease-causing or contributing variables and make up a big portion of the health claims investigation. The problem with epidemiological studies is that you cannot, except in a few instances, do clinically controlled experiments or the better of the bunch,

the Clinical Study. one test group gets a particular treatment and another control group does not. The groups are supposed to be as nearly identical as possible (same disease, same backgrounds, same ages). If there is a pronounced effect (good or bad) eureka it's a story. But sometimes the results aren't completely clear (see No.3, evaluating a study). Usually such trials are done in three phases, under FDA guidelines: Phase I: small, usually no more than a handful of people, designed strictly to look at whether the drug will cause serious complications, not whether it works. Phase II: A larger version of Phase I, again designed to look at safety, but also looking at efficacy. Phase III: Considered the definitive look at whether a new drug has merit, involving thousands of people at a number of hospitals around the country. It's usually the last step before a company asks the FDA for approval. Often the phases blur, with reports of earlier phases coming out while later phases are already in progress. and the other types:

Case-control studies depend on recollections and when dealing with mortality, of a lot of second-hand information.

Case-control studies are weaker than cohort studies.

Cross-sectional studies. Rather than dealing with individuals, these studies look at groups of one neighborhood vs. another neighborhood, blacks vs. whites, etc. These studies are very crude. They are used to indicate areas that might bear more detailed study. That's it. Cross-sectional studies are the weakest of all epidemiological studies. Meta-analysis. Some studies aren't original research but rather a review of all previous studies on the topic. Researchers draw on all the findings and try to reach some overall conclusion. Then there are Papers published in Nature, JAMA, New England Journal of Medicine and Conferences and Press Releases ... just about covers the gamut of the study arena! THE MOST DEFINITIVE QUALITY OF A STUDY While some studies may merit some specific confidence numbers (as in political polls), more often we should give any level of readership a clear understanding of just how weak or strong the study is i.e., was this "preliminary" with more testing needed; was a mild "association" strong evidence, or whatever. a. Confidence intervals: this is a measure of how confident the researchers are in their risk numbers. If the confidence interval is big, that means the numbers are not very good. Statisticians like to express their confidence in their numbers as a percentage, i.e., they are 95 percent confident that the study result is between X and U.

The trick is knowing what X and U are. For example, researchers estimated that chlorinated water was associated with a 38 percent increase in bladder cancer. Sounds like a good story? That confidence interval was 1.01 to 1.87, a wide spread. That meant that while the researchers thought it was 38 percent, it could be anywhere from 1 percent to 87 percent. Obviously, this is a highly speculative study and not worth front-page news. b. Size of effect. This is particularly important in epidemiological

studies. How much effect does the suspected agent cause? These are expressed in numbers such as .05 or 1 or 2.1. The number 1 represents a 100 percent increase in effect. That may sound big, but statisticians might still consider that insignificant.

For example, in the overall population of women, there are 10 cases per 1,000 among women who smoke. That is an effect of 1. Generally, epidemiologists like to see an effect of 3 to find significance. Sometimes 2 is okay. But usually, anything less than 1 is considered suspect, considering the weakness of the studies in the first place. In some studies this is expressed as a probability, i.e., how probable was it that the result was a fluke? It is expressed as a number from 0 to 1. If the P value is zero, then there is no chance that it was a fluke. A P value of more than 0.05 is considered weak. NOW YOU CAN BE CONFIDENT YOU'RE READING A VALID STUDY!!